Analysis of the IBD5 locus and potential gene-gene interactions in Crohn’s disease

K Negoro, D P B McGovern, Y Kinouchi, S Takahashi, N J Lench, T Shimosegawa, A Carey, L R Cardon, D P Jewell, D A van Heel

BACKGROUND AND AIMS: Genetic variation in the chromosome 5q31 cytokine cluster (IBD5 risk haplotype) has been associated with Crohn’s disease (CD) in a Canadian population. We studied the IBD5 risk haplotype in both British and Japanese cohorts. Disease associations have also been reported for CARD15/NOD2 and TNF variants. Complex interactions between susceptibility loci have been shown in animal models, and we tested for potential gene-gene interactions between the three CD associated loci.

METHODS: Family based association analyses were performed in 457 British families (252 ulcerative colitis, 282 CD trios) genotyped for the IBD5 haplotype, common CARD15, and TNF—B57 variants. To test for possible epistatic interactions between variants, transmission disequilibrium test analyses were further stratified by genotype at other loci, and novel log linear analyses were performed using the haplotype relative risk model. Case control association analyses were performed in 178 Japanese CD patients and 156 healthy controls genotyped for the IBD5 haplotype.

RESULTS: The IBD5 haplotype was associated with CD (p=0.007), but not with UC, in the British Caucasian population. The CARD15 variants and IBD5 haplotype showed additive main effects, and in particular no evidence for epistatic interactions was found. Variants from the IBD5 haplotype were extremely rare in the Japanese.

CONCLUSIONS: The IBD5 risk haplotype is associated with British CD. Genetic variants predisposing to CD show heterogeneity and population specific differences.

MATERIALS AND METHODS
Subjects
Northern European Caucasian families with one or more children affected with IBD were ascertained through the Oxford Gastroenterology Unit. Both parents were available from 457 families, of which 101 had two or more affected sibling pairs and 356 a single affected offspring. Data were obtained for a

Abbreviations: CD, Crohn’s disease; UC, ulcerative colitis; IBD, inflammatory bowel disease; TDT, transmission disequilibrium test; TNF, tumour necrosis factor α; NFκB, nuclear factor κB; PCR, polymerase chain reaction; SNP, single nucleotide polymorphism.

www.gutjnl.com
DNA). haplotype, and were essentially identical in their genetic previously reported that had alleles unique to the 250 kb

Eleven single nucleotide polymorphisms (SNPs) were previ-

Genotyping was obtained from both British and Japanese institutions.

histological findings according to conventional criteria.

healthy controls.

recruited 156 Japanese volunteers living in the same area as

Figure 1 Agarose gel image of IGR2060a_1 genotyping. First lane is 100 bp DNA size marker, followed by products of 24 individual polymerase chain reactions (PCR). PCR product sizes are control [212 bp], C allele [149 bp], and G allele [103 bp]. Arrows mark, in order, CC homozygote, GC heterozygote, GG homo-

zogygote. Lanes marked “h” are negative controls [containing no DNA].

total of 282 CD trios, 252 UC trios, and 10 indeterminate colli-

tis trios. Numbers of trios (fully genotyped for all markers) analysed after genetic stratification of the CD phenotype were:

CD-CARD15-0: 184 trios; CD-CARD15-1: 2.1: 95 trios; CD-

IBD5-0: 77 trios; CD-IBD5-1: 2.2: 200 trios; CD-TNF

exon: 251 trios; and CD-TNF

exon: 27 trios.

Japanese samples were obtained from 178 CD patients who attended Tohoku University Hospital (Sendai, Japan). We recruited 156 Japanese volunteers living in the same area as healthy controls.

In both populations, the diagnosis of IBD was made on the basis of clinical symptoms and endoscopic, radiographic, and histological findings according to conventional criteria. Demographic and phenotype data of affected CD individuals in both populations are described in table 1. Ethics approval was obtained from both British and Japanese institutions.

Genotyping

Eleven single nucleotide polymorphisms (SNPs) were previ-

ously reported that had alleles unique to the 250 kb IBDS risk haplotype, and were essentially identical in their genetic information by virtue of being in nearly total linkage disequi-

librium with one another and of similar allele frequencies.10 We genotyped two SNPs from the 5’ and 3’ ends of the IBDS haplotype (http://www.genome.wi.mit.edu/humgen/IBDS) to confirm a similar extent of linkage disequilibrium in our populations, and enable subsequent accurate identification of the IBDS risk haplotype. Genotyping for IGR2060a_1 and IGR3096a_1 was performed separately using multiplexed sequence specific polymerase chain reaction (PCR) amplifica-

tion (AmpliTag Gold, Applied Biosystems standard protocol). For IGR2060a_1, we used primers flanking the polymorphism and specific to each allele at the most 3’ base (for G allele, 5’-CATTACATCCTTGACACCTG-3’, for C allele, 5’-AGCTAGCTACATCTTGGC-3’), with the outer primers (5’-AATGTTGGAGGAAGTGTGG-3’ and 5’-TGTGAAATGGGAAATACAGT-3’) to form two allele specific PCR products of different size (fig 1). For IGR3096a_1, allele specific primers were: T allele (5’-GAACCCCAAACA TCCGGGAAAF3’), C allele (5’-CCTTGTAGTTGCTCAGCTG-

3’), and outer primers (5’-GGACAAAAATAGGCCCCACAG-3’ and 5’-TGCCACCTCCACCTCTAAG-3’). For both SNPs, the two outermost primers also provided a larger positive control band (not allele specific). PCR products were electrophoresed on agarose gels and visualised with ethidium bromide staining (fig 1).

Genotyping for common CARD15 variants (Arg702Thr, Gly908Arg, and Leu1007InsC) and the TNF’ promoter –857C/T variant was performed by PCR restriction fragment length polymorphism assay, as described previously.11

Statistical analysis

Linkage disequilibrium (D’) between variants was calculated by the ldmax algorithm implemented in the GOLD software package.12 We carried out family based association analyses (transmission disequilibrium test (TDT))13 using the ASPEx program (ftp://lahmed.stanford.edu/pub/aspx). Simulations were done to calculate empirical probabilities for the TDT χ² statistic by permuting parent alleles while fixing the identity by descent status of siblings within a family. Thus the p value for the TDT reflects an association independent of linkage when multiple siblings from the same family are tested, allowing the use of both simplex and multiply affected families in the TDT (simplex and multiply affected families were combined for all analyses). We also performed further analyses stratified on carriage of the associated disease associated mutations by recoding CD affection status as unaffected or affected (transmitted/untransmitted alleles to a single parent) from the family data (using the haplotype relative risk model of transmitted/untransmitted alleles to a single parent).14

RESULTS

Association studies of the IBDS haplotype in the British Caucasian population.

Strong linkage disequilibrium between variants from the 5’ and 3’ extremes of the IBDS haplotype (IGR2060a_1 and
families are shown in table 2. The association of CD with
sequent genotyping was therefore performed using
in the British population as found in a Canadian population.
firm a similar (minimum) extent of linkage disequilibrium
′ =0.81, p<0.00001). These data con-
D
IGR3096a_1 was observed in 96 genotyped healthy Cauca-
sian blood donors (D
IBD5 association appeared to be CD specific.
Genotype stratified association studies of CARD15, IBD5, and
variant in Caucasian CD
Table 4 Log linear analysis of variants in British Crohn’s disease. Single locus model: allele frequencies and tests of association
<table>
<thead>
<tr>
<th>Variant</th>
<th>Allele frequency</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control</td>
<td>Case</td>
</tr>
<tr>
<td>IBD5</td>
<td>39.3%</td>
<td>48.2%</td>
</tr>
<tr>
<td>CARD15</td>
<td>5.4%</td>
<td>9.8%</td>
</tr>
<tr>
<td>TNF</td>
<td>93.3%</td>
<td>95.0%</td>
</tr>
</tbody>
</table>

IGR3096a_1 was observed in 96 genotyped healthy Caucasian blood donors (D′=0.81, p<0.00001). These data confirmed a similar (minimum) extent of linkage disequilibrium in the British population as found in a Canadian population. Subsequent genotyping was therefore performed using IGR2060a_1 only.

TDT results for IBD, CD, and UC phenotypes in the British families are shown in table 2. The association of CD with IBD5 was replicated. The strength of association, as measured by the transmitted/untransmitted allele ratio for IGR2060a_1, was weaker in the British cohort (1.42) than reported in 139 Canadian trios (2.38, χ²=4.8, p=0.03). No evidence for association of IBD5 with UC was observed, and this locus appeared to be CD specific.

The weaker association of the IBD5 haplotype in the British (mean age at diagnosis of affected offspring 22.4 years) versus Canadian (16.3 years) CD affected offspring raised the possibility that the IBD5 haplotype may have an effect on age at diagnosis of CD. To test this hypothesis, we performed Kaplan-Meier analysis of the effect of IBD5 haplotype status on age at diagnosis in all CD individuals from the British cohort. IBD5 haplotype status did not have a significant effect on age at diagnosis in log rank comparisons between groups (data not shown). Mean age at diagnosis in individuals not possessing the IBD5 risk haplotype was 26.5 (SEM 1.0), heterozygotes 25.4 (SEM 0.7), and homozygotes 25.2 (SEM 0.8) years.

Genotype stratified association studies of CARD15, IBD5, and TNF, in Caucasian CD
Stratified TDT results for the CD phenotype in the British families are shown in table 3. The association, in this cohort, of the Arg702Trp and Leu1007fsinsC CARD15 variants with CD, and of TNF with CD in individuals not possessing CARD15 variants, has been reported elsewhere. The association between CD and the IBD5 haplotype was strongest in the CARD15 negative group, between CD and Leu1007fsinsC equally strong in both groups after IBD5 stratification, and between CD and TNF only in the CARD15 negative group. The association between CD and Arg702Trp was strongest in the IBD5 haplotype possession group. Transmission/non-transmission ratios (a statistic related to genotype relative risk) between groups possessing and not possessing variants were similar, and simple χ² tests of allelic transmission between groups were not significant after any stratification (table 3 and data not shown). It is noteworthy that the TDT only uses a proportion of the dataset (because only allelic transmissions from heterozygous parents are counted), and is not therefore a powerful method to test for possible epistatic interactions.

Numbers of CD trios in the TNF stratified group were small and the results from this analysis need to be interpreted with caution.

Log linear analyses of potential interactions between variants in British Crohn’s disease
The most appropriate statistical method to assess the significance of potential interactions between variants using family based data has not been established. We therefore reduced the family data to case and control alleles (sample size equivalent to a study of 230 case and 230 control individuals). We performed a log linear analysis using the haplotype relative
risk association statistic. Single locus models were first tested in which only the effect of a single genetic variant on phenotype (that is, presence of absence of CD) was analysed. Two (and similarly three) locus models allowed the main effects of multiple genetic variants to be analysed, with the addition of a higher order term(s) in the model to account for possible non-additive interactions between variants (epistasis). These models therefore allowed testing for genetic heterogeneity (that is, association still present after accounting for all non-additive interactions) and epistasis.

The Gly908Arg CARD15 variant did not show a significant association with CD, and was too rare in our population to test for interactions. As before, no association between non-stratified CD and TNF-857C was observed.

The IBD5 haplotype and the Arg702Trp/Leu1007fsinsC CARD15 variants were associated with CD (Table 4), as observed in the TDT analysis. Analysis of the main effects was then performed using two locus models, in which association was tested after conditioning for higher order effects from potential gene-gene interactions (tables 5, 6). Significant association with CD remained for the Arg702Trp and Leu1007fsinsC CARD15 variants, and the IBD5 haplotype after conditioning, providing direct evidence for independent main effects (heterogeneity). Finally, no evidence for gene-gene interactions (epistasis) were noted in two or three locus models, and the IBD5 haplotype, and the IBD5 haplotype after conditioning, providing direct evidence for independent main effects (heterogeneity).

Association studies of IBD5 variants in the Japanese population

In the Japanese population, three healthy controls and two Crohn’s disease cases were heterozygous for IGR2060a_1, and IGR3096a_1 was not polymorphic in either Japanese cohort (all wild-type). No individuals were homozygous for the IBD5 haplotype. The rarity of these variants (<1% allele frequency) in the Japanese precluded tests for association.

DISCUSSION

Much progress has been made in our understanding of the genetic susceptibility to IBD. Linkage studies have identified regions of the human genome likely to contain susceptibility loci, and more recently disease associated variants have been reported from three distinct loci (CARD15/NOD2, chromosome 5q31 cytokine cluster/IBD5 risk haplotype, and TNF). The association between CD and CARD15 variants has now been widely confirmed in Caucasian populations, and replication studies of the other two loci are awaited.

We sought first to study genetic variants from the 5q31 cytokine gene cluster in a British IBD cohort. This region contains a cluster of immunoregulatory cytokines potentially important in Th1/Th2 differentiation. However, the presence of strong linkage disequilibrium means the genetic approach alone has so far been unable to pinpoint the disease causing mutation. We confirmed the association of the IBD5 risk haplotype with CD previously reported in Canadian families. Evidence for both linkage and association in the Canadian cohort was greatest in early onset CD, and mean age at diagnosis was 4.1 years younger in the Canadian cohort than in the current British cohort. However, no effect of the IBD5 haplotype on age at diagnosis was observed in the British

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Log linear analysis of variants in British Crohn’s disease. Two locus model: allele frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohort</td>
<td>Frequencies of observed allelic combinations</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>37.9%</td>
</tr>
<tr>
<td>Case</td>
<td>42.1%</td>
</tr>
<tr>
<td>Leu1007fsinsC</td>
<td>1007fsinsC</td>
</tr>
<tr>
<td>Control</td>
<td>59.1%</td>
</tr>
<tr>
<td>Case</td>
<td>45.3%</td>
</tr>
<tr>
<td>Control</td>
<td>4.3%</td>
</tr>
<tr>
<td>Case</td>
<td>5.9%</td>
</tr>
<tr>
<td>Control</td>
<td>2.8%</td>
</tr>
<tr>
<td>Case</td>
<td>4.4%</td>
</tr>
</tbody>
</table>

Table 6 | Log linear analysis of variants in British Crohn’s disease. Two locus model: tests of association after conditioning on potential gene-gene interactions |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Crohn’s disease association (main effects)</td>
<td>χ²</td>
</tr>
<tr>
<td>IBD5, conditioned on</td>
<td>6.9</td>
</tr>
<tr>
<td>CARD15 [Arg702Trp]</td>
<td>6.3</td>
</tr>
<tr>
<td>TNF-857G</td>
<td>7.1</td>
</tr>
<tr>
<td>CARD15 [Arg702Trp], conditioned on</td>
<td>4.5</td>
</tr>
<tr>
<td>IBD5</td>
<td>6.1</td>
</tr>
<tr>
<td>CARD15 [Leu1007fsinsC], conditioned on</td>
<td>9.0</td>
</tr>
<tr>
<td>IBD5</td>
<td>15.2</td>
</tr>
<tr>
<td>TNF-857G, conditioned on</td>
<td>1.7</td>
</tr>
<tr>
<td>IBD5</td>
<td>1.4</td>
</tr>
<tr>
<td>CARD15 [Arg702Trp]</td>
<td>1.5</td>
</tr>
</tbody>
</table>
IBD5 locus and Crohn’s disease

cohort. Other differences in phenotype (for example, disease location or behaviour) might explain the stronger association in the Canadian cohort, although these comparisons could not be made as this information was not provided in the published Canadian study.17

The role of the IBD5 risk haplotype in UC has not previously been investigated. No evidence for association was observed in our British cohort, adding to the molecular evidence for heterogeneity between CD and UC obtained from studies of CARD15.

We genotyped IBD5 haplotype variants in Japanese CD and healthy controls, and found that in contrast with the Caucasian population, the variants were extremely rare. This difference is too large to be explained by the minor demographic and phenotypic differences between the populations. The two variants tested are from extreme ends of the “Caucasian” CD risk haplotype, and it therefore seems likely that the disease causing variant carried on the “Caucasian” haplotypic background does not play a role in Japanese CD (although until the variant is known this cannot be directly tested). It is also possible that other Japanese specific disease-causing mutations have arisen independently in the same gene that carries the Caucasian specific mutation. A comprehensive study is now necessary to determine whether there are variants in this region specifically associated with Japanese CD, and to assess the population specific pattern of linkage disequilibrium in the 5q31 region. If CD associated variants identified in the Japanese, this may provide an alternative method for functional studies for identifying the disease causing Caucasian IBD5 mutation. It is interesting that common Caucasian CARD15 variants have also been found to be rare in both Japanese and African American populations.22,23 It thus seems likely that different genetic defects contribute to CD susceptibility in different ethnic populations.

Epistasis, or non-additive interaction between disease causing genetic variants, is well recognised in Drosophila and mouse models,22,24 and has been demonstrated more recently in several human diseases. In IBD, a genome wide search for colitis susceptibility loci in the interleukin 10 deficient mouse demonstrated complex epistatic interactions between loci.19,25 Conversely, it is possible to knockout, over express, or mutate distinct genes involved in diverse pathways and produce a similar colitis phenotype in mouse models. Now that associated variants have been determined from three separate loci in human IBD, it is possible to assess whether epistasis plays a major role or whether heterogeneity exists between the variants. No significant interactions were found between the CARD15 variants, TNF-308, and the IBD5 haplotype in either stratified TDT or log linear analyses. This finding is supported by the lack of interaction between CARD15 variants and the IBD5 haplotype reported in a smaller Canadian cohort,17 and the IBD5 haplotype after conditioning on possible interactions between variants. Different interactions may however be present between other loci, and it seems likely that both genetic heterogeneity and epistasis may coexist in human IBD. Indeed, preliminary evidence for epistasis has already been reported between the I3p6 and IBD1 loci.16

We conclude that the IBD5 cytokine cluster risk haplotype does not influence susceptibility to UC and plays a lesser role in genetic susceptibility to CD in the British than in the Caucasian population. Our data support the existence of genetic heterogeneity and population specific differences in the inherited susceptibility to human CD. Further studies are now needed to address the relative importance of IBD susceptibility genes in different populations and in different disease subphenotypes. The possibility of epistatic interactions in IBD will need to be reassessed if associated variants are identified from other susceptibility loci.

ACKNOWLEDGEMENTS

We thank Joseph Tewari for advice on the heterogeneity analysis. This work was supported by the British Medical Research Council (DAvH, DPMcG, LINK Grant), the Wellcome Trust (LRC), the Juntendo Symposium (KN), Oxagent Ltd, and NIH grant EV13652 (LRC). We thank all participating families, our research nurses, the National Association for Crohn’s and Colitis, consultants, and general practitioners who helped ascertain families.

Authors’ affiliations

N J Lench, A Carey, Oxagent Ltd, Abingdon, UK
D P Jewell, Gastroenterology Unit, University of Oxford, UK

*K Negro and D A van Heel contributed equally to this work.

REFERENCES

13 Žubenok OS, Hughes HB 3rd, Stiffler JS. D10S1423 identifies a marker that maps to the susceptibility locus for Alzheimer’s disease (IBD5) in the mouse. Genomics 2001;77:28
Analysis of the IBD5 locus and potential gene-gene interactions in Crohn's disease

K Negoro, D P B McGovern, Y Kinouchi, S Takahashi, N J Lench, T Shimosegawa, A Carey, L R Cardon, D P Jewell and D A van Heel

Gut 2003 52: 541-546
doi: 10.1136/gut.52.4.541

Updated information and services can be found at:
http://gut.bmj.com/content/52/4/541

These include:

References
This article cites 29 articles, 6 of which you can access for free at:
http://gut.bmj.com/content/52/4/541#BIBL

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections
Ulcerative colitis (1113)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/